

(19)



Europäisches Patentamt
European Patent Office
Office européen des brevets



(11) Publication number:

0 680 237 A2

(12)

EUROPEAN PATENT APPLICATION(21) Application number: **95102738.2**(51) Int. Cl.⁶: **H04Q 11/04, H04L 12/56**(22) Date of filing: **27.02.95**(30) Priority: **23.03.94 GB 9405788**(43) Date of publication of application:
02.11.95 Bulletin 95/44(84) Designated Contracting States:
AT BE CH DE DK ES FR IT LI SE

(71) Applicant: **ROKE MANOR RESEARCH LIMITED**
Roke Manor
Romsey,
Hampshire, SO51 0ZN (GB)
Applicant: **Siemens Aktiengesellschaft**
Postfach 22 16 34
D-80506 München (DE)

(72) Inventor: **Hayter, Andrew Timothy**
3 Cobden Gardens
Blitstone Park,
Southampton SO2 4LN (GB)

Inventor: **Momtahan, Paul Parsheed**
10 Hunters Crescent
Totton,
Southampton SO4 2FA (GB)
Inventor: **Davis, Simon Paul**
17 Westering
Romsey,
Hampshire SO51 7LX (GB)
Inventor: **Wallmeier, Eugen Bernhard**
Bussardstrasse 14
D-82223 Eichenau (DE)

(74) Representative: **Allen, Derek**
Siemens Group Services Limited
Intellectual Property Department
Roke Manor
Old Salisbury Lane
Romsey
Hampshire SO51 0ZN (GB)

(54) **ATM queuing and scheduling apparatus.**

(57) The apparatus comprises a buffer store connected to an input line upon which ATM cells are received, and an output line upon which ATM cells are transmitted. A control means receives, for each cell, a channel identifier and a path identifier from which a first signal is generated indicative of a maximum delay value which is applied to the buffer store and used to schedule the particular cell for transmission, on a calendar controlling a sustainable cell rate. The control means generates a second signal indicative of a minimum delay value which is applied to the buffer means and, in conjunction with said first signal is used to schedule the particular cell for transmission, on a calendar controlling the peak cell rate. The apparatus can also handle routing information, where a group of channel connections define a Route which share a peak cell rate for the Route and are scheduled on the peak cell rate calendar.

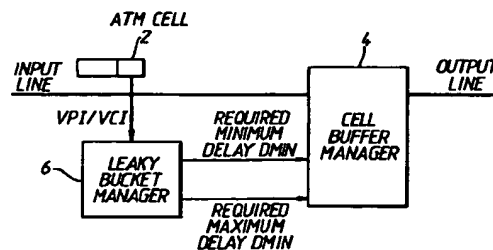


Fig. 4

EP 0 680 237 A2

The present invention relates to apparatus for queuing and scheduling ATM cells in an ATM switch.

ATM networks and switches support a mixture of traffic including bursty traffic. By its nature, bursty traffic requires high bit rates for part of the time and little or no bit rate for the rest of the time. In order to use efficiently the bit rate available in a network, it is necessary to allocate a connection a lower bit rate than its peak bit rate such that the total peak bit rate of all the connections may be greater than the bit rate of the network links.

To support this situation an ATM multiplexer requires large buffers at the input of the switch to buffer cells when they arrive at the multiplexer and cannot be transmitted across it and also buffers at the output of the switch so that traffic may be shaped, whereby the profile of the traffic is altered before transmission from the switch. Such a switch with large buffers is termed a large buffered switch.

In an ATM network, a user, who wishes to set up a connection orientated virtual channel connection in a quality of service category where statistical gain is required, will have to negotiate with the network at least three traffic parameters. These parameters will include the peak cell rate, the sustainable cell rate and some measure of the maximum burst size. Output buffers are therefore required to store cells so that they are transmitted at their sustainable cell rate but not above their peak cell rate. This process must be completed for a plurality of ATM virtual channel connections.

In German patent application number 93120828.4, an egress shaping algorithm is discussed.

If the sum of the sustainable cell rates for connections accepted for an egress network link can be handled by the bit rate available for statistically multiplexed services on a network link, the egress scheduler/shaper disclosed in the application referred to above can guarantee a minimum service rate equal to the sustainable cell rates for each connection on a link using a simple resource allocation procedure which can provide high levels of statistical gain with bounded buffer sizes.

Referring to Figure 1, this shows how ATM cells 2 enter the system on the input line, this line will have come from the part of the device which has multiplexed all the cells destined for this egress queue. The virtual path identifier/virtual channel identifier VPI/VCI of the cell is used to obtain the maximum allowed delay D_{max} in order to maintain the sustainable cell rate as the minimum service rate for that virtual channel connection. The value D_{max} is then fed into the cell buffer manager 4 from the leaky bucket manager 6. The cell buffer manager 4 schedules the cell for transmission via a calendar mechanism. The output line

transfers cells from the calendar, however, they may be multiplexed together with other cells before being transmitted onto the network egress link.

The calendar mechanism used is shown in Figure 2, in which the following abbreviations are used. FLT is the Free List Tail, FLH is the Free List Head, CET is the Calendar Entry Tail, CEH is the Calendar Entry Head, OLT is the Output List Tail and OLH is the Output List Head. The constituent components of the cell buffer manager are the cell memory, the calendar and a series of pointers. When a cell arrives at the cell buffer manager, a free cell location is taken from the head of the free cell list and the cell is copied to this location. The cell is then scheduled onto the calendar at time $T + D_{max}$ (modulo N_{CA}) where time T is the current time and N_{CA} is the size of the calendar. If another cell is scheduled for the same time slot, then it is added to the front of the queue within the calendar time slot.

There are two output time pointers, a real time pointer T , and a read pointer RP which can move ahead of the real time pointer. While the RP and T pointers point to the same time slot, the complete linked list of cells connected to a time slot are transferred to the output queue. If the RP pointer moves ahead of the T pointer, then when it finds a time slot with cells in it, the cells are moved to the output queue one at a time so that there is never more than one cell in the output buffer. The reason for this is that the RP pointer is moving ahead of the time when the cell must be transmitted from the output calendar, denoted by pointer T . If a cell arrives at the output calendar and it requires to be scheduled immediately at time T , the cell must be transferred to the output queue directly, and should be transmitted from the output queue without delay. For this reason the output queue must never have more than one cell in it when the pointer RP is ahead of the real time pointer T . However, in this situation, pointer RP may remain at its position to a cell slot ahead of pointer T .

If a cell arrives at the calendar and requires to be scheduled between pointers T and RP , the cell is scheduled, but also the pointer RP is returned to the time slot position of this newly scheduled cell. After processing it, the pointer RP then searches through the calendar again for non-empty time slots. The reason for this is that if the cell arrival rate is greater than the transmission rate, many cells may arrive at the calendar and require scheduling between pointers RP and T . Simply transferring them to the output port will leave cells with later departure times being transmitted before cells with immediate departure times. The pointer RP must always be pointing to the closest time slot to pointer T , with cells in it.

In the situation where the pointer RP is ahead of the pointer T, if the load on the port increases, the linked lists associated with each time slot will increase in length and eventually the real time pointer T will 'catch up' with the read pointer RP.

Thus it is possible for the read pointer to get ahead of the current time T but never behind. Hence this algorithm will guarantee the sustainable cell rate for a given virtual channel connection but has no limit on the peak cell rate.

Therefore, an aim of the present invention is to provide apparatus for limiting the peak cell rate.

According to the present invention there is provided apparatus for queuing and scheduling ATM cells connected to an output side of an ATM switch comprising buffer means connected to an input line upon which ATM cells are received, an output line upon which ATM cells are transmitted, control means for receiving, for each cell, a channel identifier and a path identifier from which a first signal is generated indicative of a maximum delay value which is applied to the buffer means and used to schedule the particular cell for transmission, on a calendar controlling a sustainable cell rate, characterised in that said control means generates a second signal indicative of a minimum delay value which is applied to the buffer means and, simultaneously with said first signal is used to schedule the particular cell for transmission.

The minimum delay value is used to control a further calendar for controlling a peak cell rate.

An embodiment of the invention will now be described with reference to Figures 3 to 7 of the accompanying drawings wherein;

FIGURE 3 shows a typical layout of an ATM switch,

FIGURE 4 shows a block diagram of an extended egress queuing and scheduling system,

FIGURE 5 shows the peak cell rate and the sustainable cell rate calendars,

FIGURE 6 shows the interaction between the peak cell rate and sustainable cell rate calendars, and

FIGURE 7 shows the peak cell rate and sustainable cell rate calendars extended to handle 'Routes' as described hereinafter.

Referring to Figure 3, on the ingress side are shown line cards 8, the outputs of which are connected to an input of an ATM multiplexer 12. The output of the multiplexer 12 is connected to an input of a statistical multiplexing unit 16 the output of which is fed to the ATM switching network 20. The output of the switching network 20 is connected to an input of a further statistical multiplexing unit 18, the output of which is connected to an input of an ATM multiplexer 14. The outputs of the ATM multiplexer 14 are connected to the input of a number of line cards 10 which are on the egress

side of the switching network. The links between the various components designated M represent multiplexed internal links. Each statistical multiplexing unit has a flow control controller 22. The combination, for example, of the line cards 8, multiplexer 12 and statistical multiplexing unit 16 comprises a peripheral switch group. In practice there will be a number of peripheral switch groups connected to the ATM switching network. Within each statistical multiplexing unit 16 there is one input queue for each of the peripheral switch groups attached to the ATM switching network. Cells may be sent independently between any of the peripheral switch groups. There are limiting factors which control the cell rate, and they are the output link bit rate from the ATM switching network to a peripheral switch group, and the output link bit rate from a peripheral switch group to the ATM switching network.

Flow control procedures operate to manage these limited bit rates fairly for all connections, both internally to the switch and between the peripheral switch groups, and to limit bit rate to peak reservation across the ATM switching network.

The present invention is conveniently located in the statistical multiplexing unit 18 and will now be described.

Referring to Figure 4, which bears the same references as Figure 1, to take account of the peak cell rate, a similar value to that of Dmax (maximum interdeparture time) must be generated by the leaky bucket manager, corresponding to the earliest time at which a cell can be transmitted. This value, Dmin is defined as the minimum delay or 'number of cell slots' before the cell can be transmitted.

The general principle of operation for this mechanism is that a cell can only be sent between the time interval calculated from its peak cell rate and the time interval calculated from its sustainable cell rate. Under low load conditions cells will be sent closer to the peak cell rate whereas at high loads this will tend towards the sustainable cell rate. The method described below uses two scheduling mechanisms for the same cell. One of these mechanisms schedules the cell for the minimum time, the peak cell rate, the other schedules the cell for the maximum time sustainable cell rate. The peak cell rate calendar is read in real time which ensures that cells can not be sent before the peak cell rate criterion is met. However, the sustainable cell rate calendar may be read before but never after the maximum time, sustainable cell rate, thus allowing a service rate of \geq sustainable cell rate. When both scheduled events (from the peak cell rate and sustainable cell rate calendars) are processed the real cell can be sent.

The two values from the leaky bucket manager 6 (Dmax and Dmin) are used by the new cell buffer manager 4 to schedule events on to two separate calendars. The first calendar termed the sustainable cell rate calendar, is exactly the same as that described in the above referenced German patent application. The second calendar is called the peak cell rate calendar.

Referring to Figure 5, a comparison is shown between the sustainable cell rate calendar and the peak cell rate calendar.

The value $T + Dmin$ is used to schedule a cell onto the calendar, and a real time read pointer RTRP is used instead of the read pointer. The pointer RTRP is incremented each time slot and thus can not run ahead of time. The pointer RTRP is also equal to the real time pointer T, used for the sustainable cell rate calendar shown in Figure 2.

In Figures 5 and 6, the following abbreviations are used. FLT is the Free List Tail, FLH is the Free List Head, CET is the Calendar Entry Tail, CEH is the Calendar Entry Head, RP is the Read Pointer, T is the Real Time Pointer, POLT is the Peak Output List Tail, POLH is the Peak Output List Head, SOLT is the Sustainable Output List Tail, SOLH is the Sustainable Output List Head and NCA is the Size of Calendar.

Each cell only needs to be stored once on the cell memory, but requires two pointers; the sustainable cell rate pointer and the peak cell rate pointer. How this works logically is illustrated in Figures 5 and 6. In Figure 5, cells a, b, d and e are shown as scheduled on both the sustainable cell rate calendar and the peak cell rate calendar. Cells b and a form a time slot queue on the peak cell rate calendar. The peak cell rate pointers for these cells are linked together and attached to a peak cell rate calendar time slot indicating their minimum departure time delay. However, cells b, d and e form a time slot queue on the sustainable cell rate calendar. The sustainable cell rate pointers for these cells are linked together and attached to a sustainable cell rate calendar time slot indicating their maximum departure time delay. The interaction between these two calendars with a common cell memory is shown in Figure 6. Referring to Figure 6, cell b is shown as linked onto both calendars via separate pointers. In this illustration the pointers and cells are shown adjacent to each other, however, this is a logical representation; a cell is stored in memory and a pointer identifies it, however, the pointer may exist in a different area of memory to the cell that it identifies.

The pointers T and RTRP point to cell slots on different calendars, but which represent the same time, the peak cell rate calendar time slot is always processed before the sustainable cell rate calendar time slot. Processing of the peak cell rate time slot,

requires all the cells attached to it to have their peak cell rate pointer values converted to a known null value, for example zero. From the sustainable cell rate calendar, all cells processed with their peak cell rate pointer value set to zero may be transferred to the output queue. Cells attached to the sustainable cell rate calendar which do not have their peak cell rate pointers set to zero when they are processed by the sustainable cell rate calendar are left on the sustainable cell rate calendar. Therefore, the peak cell rate calendar time slot must always be processed first, as explained above to minimise cells left on the calendar.

The cells left on the sustainable cell rate calendar which have not had their peak cell rate pointers set to zero are monitored such that when their peak cell rate pointers are set to zero, the sustainable cell rate read pointer RP moves back to them and transfers them to the output queue. From this new point it then starts searching forward again.

The invention as described above may be extended to 'Routes'. Routes are defined for a group of virtual channel connections with separate quality of services but the same peak cell rate. Connections accepted on the basis of their sustainable cell rate may use Routes to define both the peak cell rate and implicitly the path to the next large buffered switch through a number of cross-connects. The calendar scheduling algorithm must thus be able to schedule cells according to their sustainable cell rate value associated with their virtual channel connection and to their peak cell rate value associated with their Route.

The diagram in Figure 7 shows the calendar in schematic form with the additions required for handling Routes.

This quite clearly handles Routes with only a minor modification to the look-up table so that the peak cell rate for a Route is shared between several virtual channel connections VCC's, and may thus be jointly scheduled by the peak cell rate calendar.

It will therefore be appreciated that in a network composed of a number of ATM switches, a bursty data connection will be set up by defining the traffic characteristics for example the sustainable cell rate, the peak cell rate and the maximum burst size. While passing through the ATM switch the traffic characteristics may alter to be outside the agreed parameters, and therefore liable to be deleted. The invention as described performs traffic shaping on VCC connections of ATM cells so that they are output from the switch within the parameters assigned to them, namely the sustainable cell rate, the peak cell rate and the maximum burst size.

It will be readily appreciated by those skilled in the art that various modifications are possible without departing from the scope of the invention as claimed.

Claims

1. Apparatus for queuing and scheduling ATM cells connected to an output side of an ATM switch comprising buffer means connected to an input line upon which ATM cells are received, an output line upon which ATM cells are transmitted, control means for receiving, for each cell, a channel identifier and a path identifier from which a first signal is generated indicative of a maximum delay value which is applied to the buffer means and used to schedule the particular cell for transmission, on a calendar controlling a sustainable cell rate, characterised in that said control means generates a second signal indicative of a minimum delay value which is applied to the buffer means and, simultaneously with said first signal is used to schedule the particular cell for transmission.
2. Apparatus as claimed in Claim 1, wherein the minimum delay value controls a further calendar for controlling a peak cell rate.
3. Apparatus as claimed in Claim 2, wherein a value $T + D_{min}$ is used to schedule a cell onto the second calendar, where T = real time and D_{min} is the minimum delay value.
4. Apparatus as claimed in Claim 3, wherein the second calendar uses a real time read pointer for each cell which is incremented for each time slot.
5. Apparatus as claimed in Claim 4, wherein each cell is stored only once in said buffer means and has a first pointer identifying its time slot in the peak cell rate calendar, and a second pointer identifying its time slot in the sustainable cell rate calendar.
6. Apparatus as claimed in Claim 5, wherein a group of channel connections defining a Route, share a peak cell rate for the Route, and are scheduled on said peak cell rate calendar as a Route but scheduled on the sustainable cell rate calendar individually.

5

10

15

20

25

30

35

40

45

50

55

5

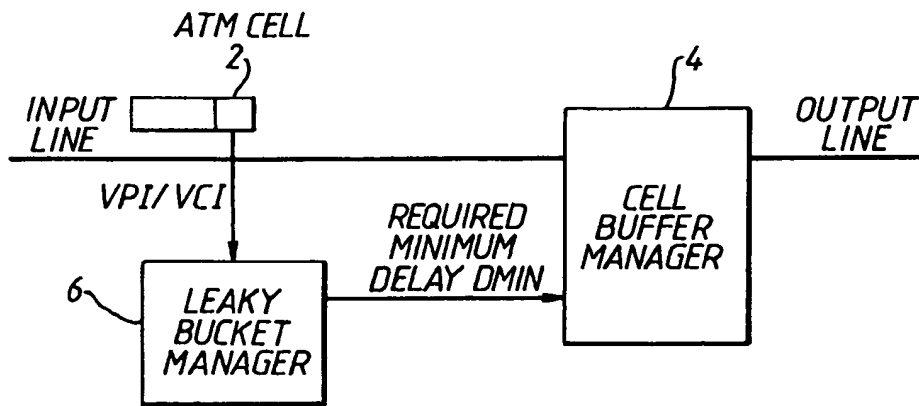


Fig. 1

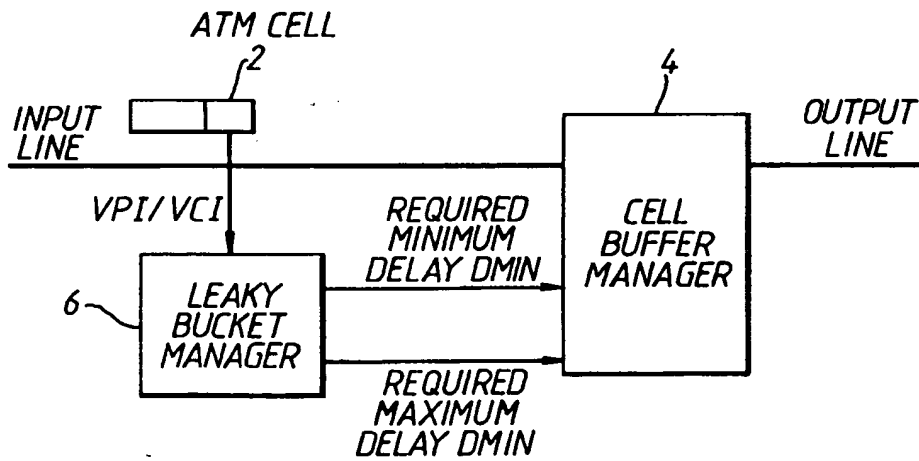


Fig. 4

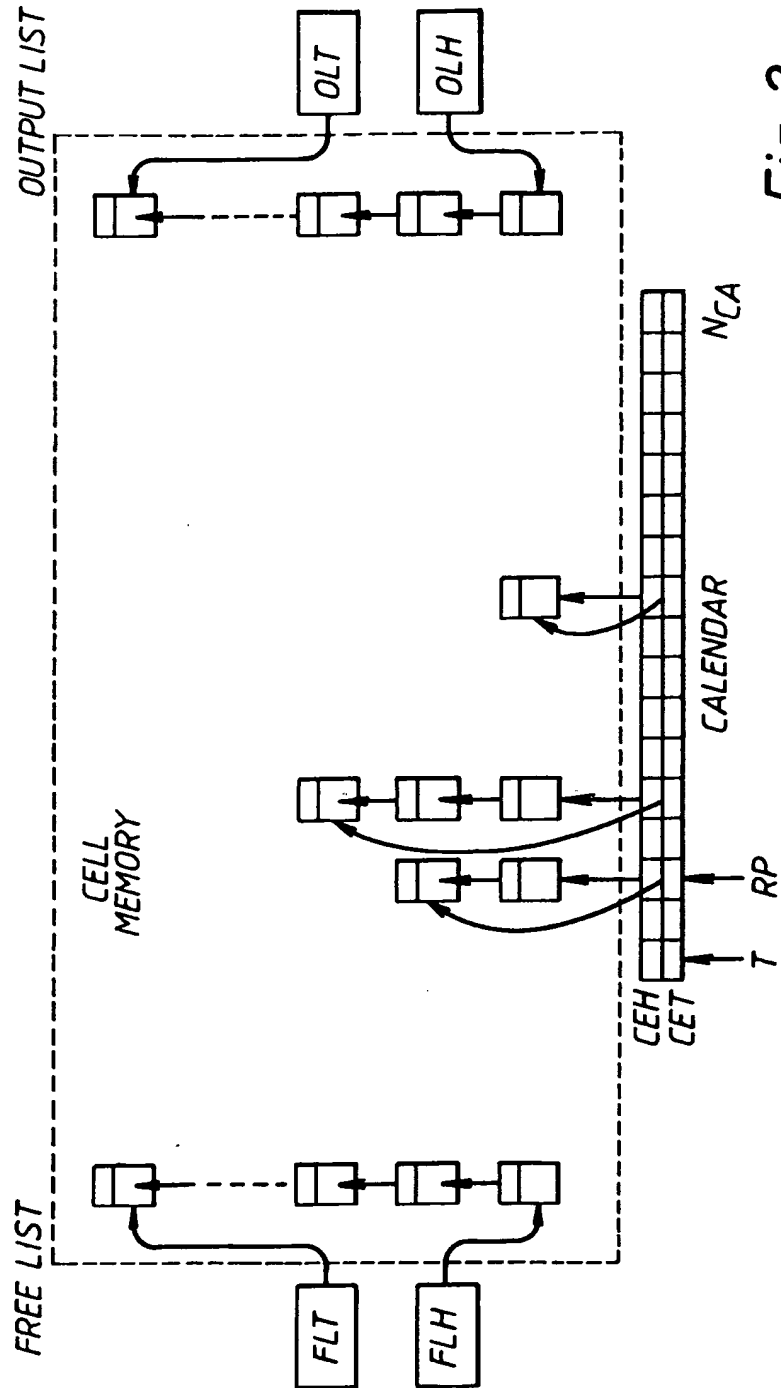


Fig. 2

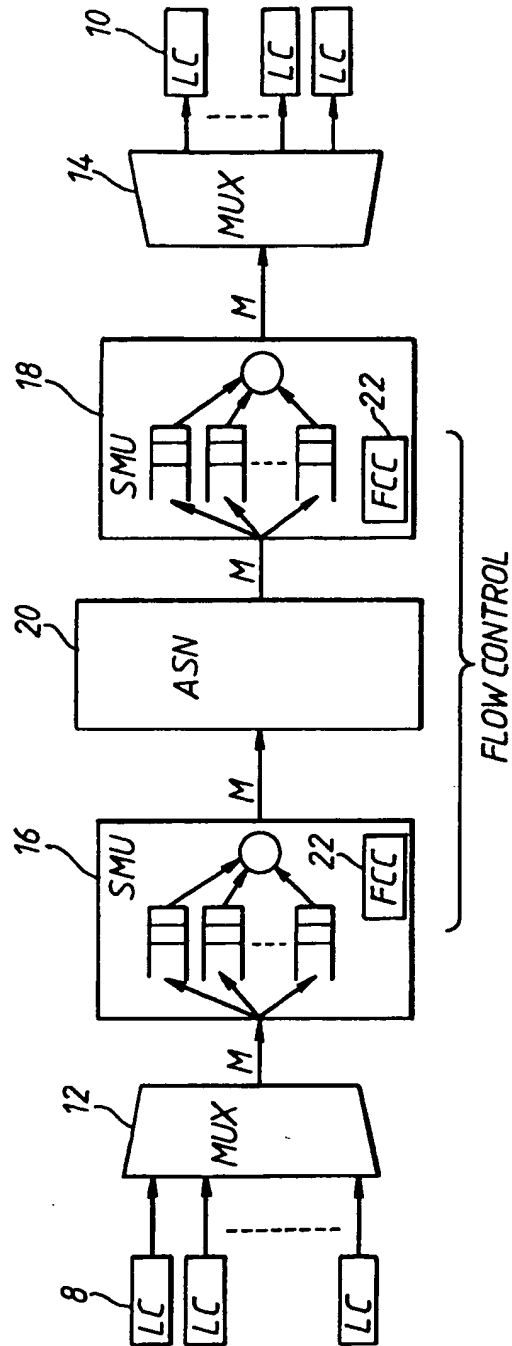


Fig. 3

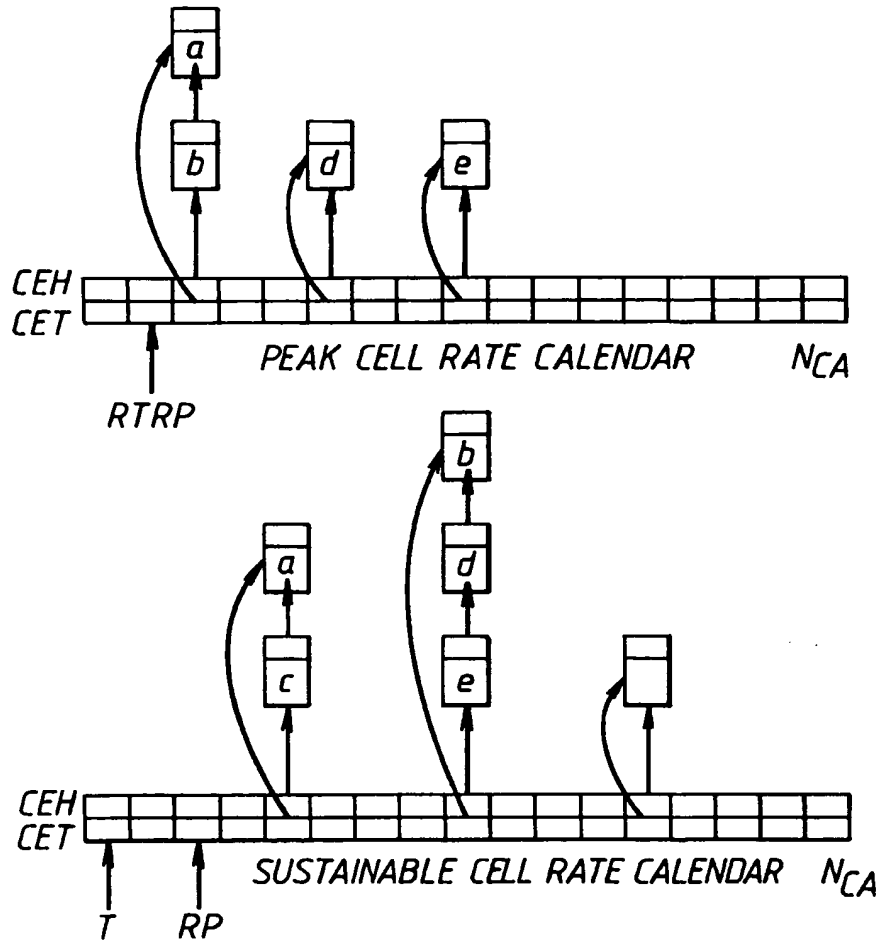


Fig. 5

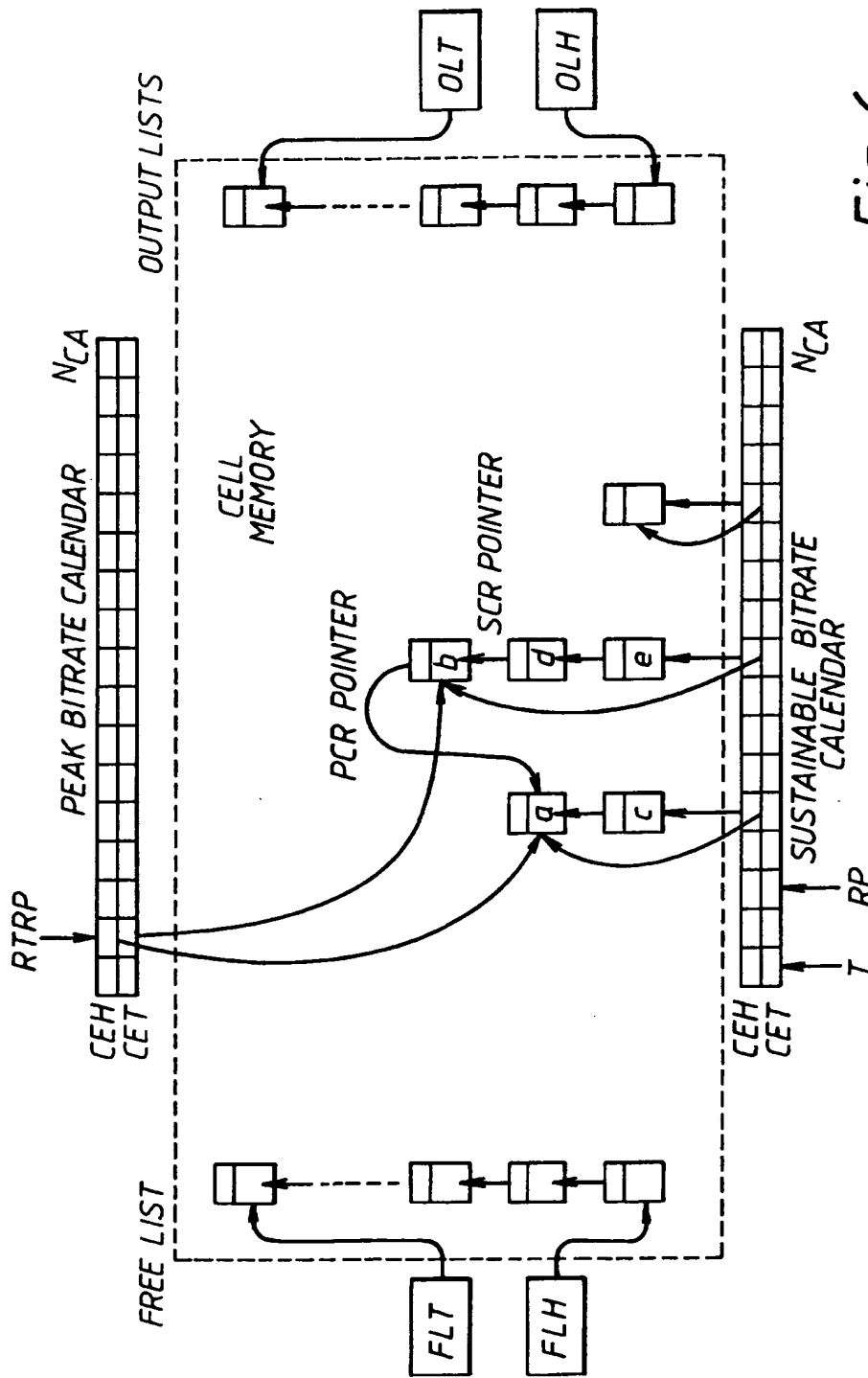


Fig. 6

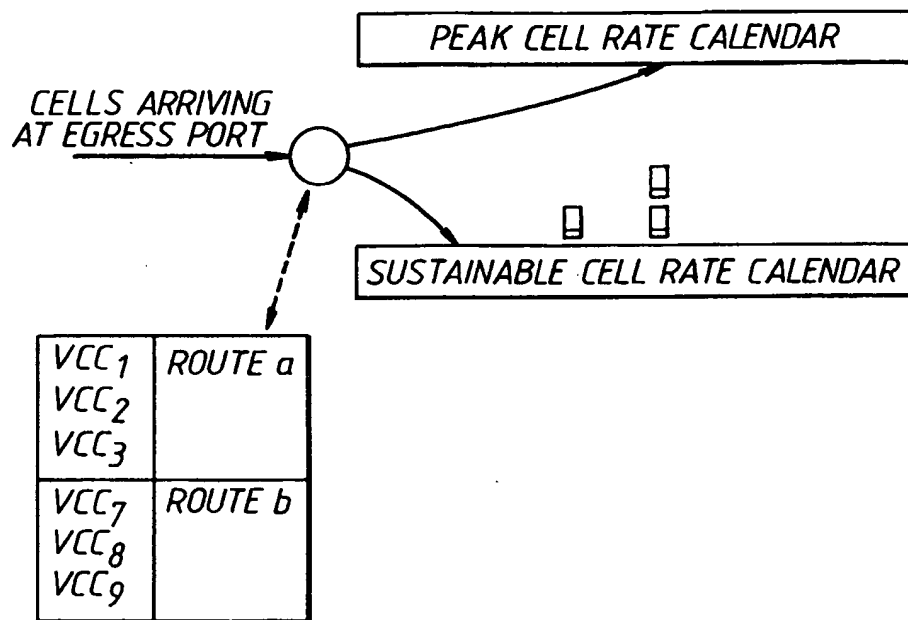


Fig. 7